

# Psychological Assessment

## Validation of the Intrinsic Spirituality Scale (ISS) With Muslims

David R. Hodge, Tarek Zidan, and Altaf Husain

Online First Publication, May 25, 2015. <http://dx.doi.org/10.1037/pas0000130>

### CITATION

Hodge, D. R., Zidan, T., & Husain, A. (2015, May 25). Validation of the Intrinsic Spirituality Scale (ISS) With Muslims. *Psychological Assessment*. Advance online publication. <http://dx.doi.org/10.1037/pas0000130>

# Validation of the Intrinsic Spirituality Scale (ISS) With Muslims

David R. Hodge  
Arizona State University and University of Pennsylvania

Tarek Zidan  
James Madison University

Altaf Husain  
Howard University

This study validates an existing spirituality measure—the intrinsic spirituality scale (ISS)—for use with Muslims in the United States. A confirmatory factor analysis was conducted with a diverse sample of self-identified Muslims ( $N = 281$ ). Validity and reliability were assessed along with criterion and concurrent validity. The measurement model fit the data well, normed  $\chi^2 = 2.50$ , CFI = 0.99, RMSEA = 0.07, and SRMR = 0.02. All 6 items that comprise the ISS demonstrated satisfactory levels of validity ( $\lambda > .70$ ) and reliability ( $R^2 > .50$ ). The Cronbach's alpha obtained with the present sample was .93. Appropriate correlations with theoretically linked constructs demonstrated criterion and concurrent validity. The results suggest the ISS is a valid measure of spirituality in clinical settings with the rapidly growing Muslim population. The ISS may, for instance, provide an efficient screening tool to identify Muslims that are particularly likely to benefit from spiritually accommodative treatments.

*Keywords:* Muslims, spirituality, Islam, cultural competence, measurement

It is increasingly recognized that spirituality can play an important role in clinical practice (Pargament, 2013). A substantial body of empirical research has documented a generally positive relationship between spirituality and a variety of salutary mental health outcomes (Koenig, King, & Carson, 2012). This research suggests that spirituality is an important asset that can often be operationalized in clinical settings to enhance wellness.

A number of instruments have been developed to measure spirituality in clinical settings (Monod et al., 2011). Most of these measures, however, have been validated with relatively homogeneous samples (Hill & Hood, 1999). The typical validation sample consists primarily of Christians, especially young university students who are Christians (Kapusinski & Masters, 2010).

This can limit the utility of spirituality instruments when assessing spirituality among non-Christian samples. Faith groups tend to affirm distinct value systems. As has been widely observed, spirituality and religion represent distinct manifestations of cultural diversity (Koenig, 1998; Richards & Bergin, 2014; Van Hook, Hugen, & Aguilar, 2001). As the United States grows increasingly religiously diverse (Eck, 2001), the importance of instruments that have been validated with other, non-Christian groups grows (Kapusinski & Masters, 2010). Accordingly, the purpose of this study is to validate an existing spirituality instrument with Muslims,

perhaps the largest non-Christian religious group in the United States.

## Literature Review

Muslims typically affirm a culturally distinct value system (Graham, Bradshaw, & Trew, 2010). Islam provides adherents with a unique way of life that unifies metaphysical and materialistic domains. As such, Islam represents a distinct worldview that differs from both Christianity and the dominant secular culture (Smith, 2003).

It is important to note that the exact parameters of the Islamic value system are contested. Numerous contextual variables influence understandings of Islam (Dwairy, 2006). For instance, nation of origin, cultural heritage, educational status, interpretative tradition, degree of acculturation to the dominant secular culture, and many other factors can shape beliefs and values at the level of the individual. Within this diversity, however, a number of tenets are widely affirmed.

At the heart of the Islamic value system is belief in God (Kobeisy, 2004). God is the omniscient Creator who, in His infinitely mercy, wants the best for His people. Accordingly, individuals are expected to carry out God's will rather their own. Indeed, the word *Islam* means submission to the will of God.

The specifics of one's submission are guided by a number of sources. The most important of these is the Quran (God's revelation to the prophet Muhammad). Other notable sources of guidance include the Hadith (sayings of the prophet), the Sunnah (actions of the prophet), and jurists' interpretations of these texts (Koenig & Shohaib, 2014).

These sources provide the basis for key Islamic tenants, such as the "five pillars" (Dwairy, 2006). These pillars consist of (a) the declaration of faith, (b) prayer at five specific times throughout the day, (c) charitable alms or giving, (d) yearly

---

David R. Hodge, School of Social Work, Arizona State University and Program for Research on Religion and Urban Civil Society, University of Pennsylvania; Tarek Zidan, Department of Social Work, James Madison University; Altaf Husain, School of Social Work, Howard University.

Correspondence concerning this article should be addressed to David R. Hodge, Mail Code 3920, CoPP, 411 North Central Avenue, Suite 800, Phoenix, AZ 85004-0689.

fasting during the month of Ramadan from sunrise to sunset, and (e) a pilgrimage to Mecca at least once in one's lifetime. These pillars form the core of a wider belief system that typically includes other values, such as the importance of community, family, and modesty (Husain & Ross-Sheriff, 2011).

Most Muslims in the United States appear to take their faith seriously. For example, according to the Pew Research Center (2011) 69% of American Muslims report that religion is "very important" in their lives. An additional 22% indicate that religion is "somewhat important" and only 8% report that religion is "not too" or "not at all important."

In light of the distinctive Islamic value system, the importance of using measures that have been validated with Muslims in the United States has been widely noted (Ghorbani, Watson, & Khan, 2007; Jana-Masri & Priester, 2007; Ji & Ibrahim, 2007; Kapuscinski & Masters, 2010). People perceive reality through their respective worldviews. These worldviews shape beliefs and practices. In turn, they influence how people interpret and respond to the questions that comprise specific measures (de Klerk, Boshoff, & Wyk, 2009).

The importance of culturally relevant measures is underscored by the size of the Muslim population in the U.S. The Pew Research Center (2011) places the number of Muslims at 2.75 million while Bagby's survey of American mosques (Bagby, 2011) suggests that approximately seven million Muslims live in the United States. In either case, it is clear the number of Muslims is substantial and growing rapidly (Pew Research Center, 2011).

In keeping with their growing presence on the American landscape, some religion measures have been validated with U.S.-based samples of Muslims (Abu-Raiya & Hill, 2014). A case in point is the Islamic Behavioral Practices scale (Jana-Masri & Priester, 2007). As the name implies, this measure taps common religious beliefs and practices within Islam.

Increasingly, however, efforts have been made to distinguish spirituality from religion (Pargament, 2013). Traditionally, spirituality and religion have been viewed as identical constructs. Indeed, many members of the general public use the two terms in an essentially interchangeable manner (Zinnbauer et al., 1997).

Contemporary scholars, however, tend to conceptualize religion and spirituality as overlapping but distinguishable entities (Kapuscinski & Masters, 2010). In brief, religion is typically understood to refer to communally held beliefs and practices that are related to the sacred or transcendent. Conversely, spirituality refers to an individual's subjective relationship with God, or a sacred or transcendent dimension of existence.

Spirituality plays a critical role in the lives of many Muslims (Hall, 2012). As implied above, submission to God is perhaps the central tenant in Islam (Kobeisy, 2004). Although individual Muslim's relationship with God can and does vary, Islam posits that God's will should guide and direct people's lives.

Despite the importance of spirituality in the lives of many Muslims, few spirituality instruments have been validated with members of this population. At least two reviews have specifically examined spirituality measures (Kapuscinski & Masters, 2010; Monod et al., 2011). None of the identified measures appeared to have been validated with Muslims.

Among the instruments discussed in these reviews, the Intrinsic Spirituality Scale (ISS) may be a particularly suitable choice for

cross-cultural validation. The ISS is designed to measure spirituality among both theistic and nontheistic populations regardless of whether or not respondents express their spirituality within or outside of religious frameworks. Based upon Allport and Ross' (1967) intrinsic conceptualization, the ISS measures the degree to which spirituality operates as an intrinsic motivational drive (Hodge, 2003). Put differently, the ISS assesses the degree to which spirituality functions as an individual's master motive, or perhaps more simply, the degree to which one's spirituality guides and directs one's life. This aligns with the Islamic tenet that one's relationship with God should guide and direct adherents' lives.

Ghorbani, Watson, and Khan (2007) have suggested that the intrinsic conceptualization has particular resonance within the Islamic tradition. These scholars posit that the intrinsic measure, at a conceptual level, is highly consistent with the Muslims' experiential connection with God. In short, their theoretical and empirical work suggests the ISS is culturally congruent with an Islamic worldview.

A preliminary examination of the ISS by the research team—all of whom have some expertise in Islamic discourse—confirmed this perspective. This assessment of the instrument's face validity was also corroborated by qualitative reports of Muslims' perspective on culturally congruent measures of spirituality and religion (Berry, Bass, Forawi, Neuman, & Abdallah, 2011). Yet, in spite of its apparent utility with Muslims, the ISS was validated with a sample of predominantly Christian university students, raising questions about its validity with Muslims (Hodge, 2003). Taken together, these various lines of conceptual, theoretical, and empirical work suggested the ISS was an ideal candidate for cross-cultural validation.

### The Present Study—Hypotheses

To validate the ISS with Muslims, a specialized form of structural equation modeling (SEM) was conducted, namely a confirmatory factor analysis (CFA). CFA is a theoretically derived statistical procedure that examines the relationship between observed indicators and latent constructs, such as spirituality, in the form of a measurement model (Kline, 2011). This statistical procedure is often used to study the psychometric properties of instruments in cross-cultural contexts.

In the current study, it was hypothesized that the measurement model would fit the data well using a sample of self-identified Muslims. In addition, it was hypothesized that each individual indicator would exhibit high levels of validity and reliability within the context of a SEM framework. Using Cronbach's alpha to assess reliability, it was also expected that high levels of internal consistency would be obtained with the study's sample.

Supplemental tests were conducted to assess criterion validity—the degree to which an instrument correlates with a relatively well established measure (Monod et al., 2011). No "gold standard" spirituality instrument currently exists. Consequently, previous validation studies have used established religion measures to assess criterion validity, such as some version of Allport and Ross' (1967) intrinsic measure of religion.

Following this practice, criterion validity was assessed using two measures: a measure of intrinsic religion (Gorsuch & McPherson, 1989) and the Islamic Behavioral Practices scale (Jana-Masri & Priester, 2007). Because most work with the intrinsic measure of

religion has been conducted with Christian majority samples, it seemed appropriate to include a religion measure that has been validated with Muslims in the United States, in addition to the intrinsic measure.

In keeping with the view that spirituality and religion are overlapping constructs, it was hypothesized that the ISS would correlate strongly with both religion measures (Ghorbani et al., 2007). Concurrently, in keeping with the notion that spirituality and religion are distinguishable entities, it was also expected that a perfect correlation would not exist. Because spirituality taps a related, but not identical construct, the correlation coefficient should be consistent with this theoretical relationship. In other words, the correlation should be strong but not approach 1.00.

Finally, tests were conducted to assess concurrent validity—the extent to which a measure correlates with other, theoretically related variables (Monod et al., 2011). To assess this form of validity, prior validation research has used various psychological constructs, such as life satisfaction and depression (Kapuscinski & Masters, 2010; Monod et al., 2011).

Accordingly, these two psychological constructs were used to assess concurrent validity. In keeping with previous research, it was posited that the ISS would be positively associated with life satisfaction and negatively associated with depression (Koenig et al., 2012). The methods used to test these three sets of hypotheses are delineated next.

## Method

### Research Design

Surveys were administered—both in person and online—at essentially a single point in time. The research was conducted with the approval of a university institutional review board (IRB). The present study was part of a larger research project designed to provide more effective, culturally relevant services to Muslims living in the United States.

### Sampling and Data Collection Procedures

To obtain a diverse sample of Muslims, a purposive snowball sampling strategy was used (Babbie, 2013). Initially, hard copies of the survey ( $n = 19$ ) were administered at local Mosques/Islamic centers. The personal administration of the surveys allowed respondents the opportunity to provide verbal feedback on the survey before it was placed online. All surveys were written in the English language and no translators were used. None of the respondents noted any problems regarding the survey instrument.

Subsequently, various Islamic organizations known to the research team were contacted. The purpose of the study was explained and permission was sought to administer the survey online to their members. In addition, individuals were asked to identify other Islamic organizations that might be willing to participate in the study.

Counting Mosques and Islamic centers, surveys were administered at 22 different organizations. The organizations were diverse in mission and size, and encompassed local, state, and national associations. Studies suggest the data obtained in paper-based and online surveys is generally similar (Barrios, Villarroya, Borrego, & Olle, 2011; Shin, Johnson, & Rao, 2012), although some research

suggests educated professionals may be more likely to participate in online surveys (Barrios et al., 2011).

### Instrument

In addition to demographic items, the self-report survey instrument included a number of items to measure the constructs mentioned in the previous section. These constructs included spirituality, intrinsic religion, Islamic practices, life satisfaction, and depression.

The ISS (Hodge, 2003), the focus of the study, is comprised of six items. Each item is formatted using the phrase completion methodology (Hodge & Gillespie, 2005). With this approach individuals are presented with a phrase (e.g., *In terms of the questions I have about life, my spirituality answers . . .*) and a response key on which to complete the phrase. The response key features a 0 to 10 continuum, in which zero represents the absence of the entity being assessed and 10 represents the theorized maximum (e.g., 0 = *no questions*, 10 = *absolutely all my questions*). Spirituality is defined in the scale's orienting material as one's relationship to God or whatever is perceived to represent Ultimate Transcendence. The Cronbach's alpha coefficient obtained in the original validation study was .96. In a subsequent validation study with a sample of Alzheimer's caregivers, an alpha of .92 was obtained (Gough, Wilks, & Prattini, 2010).

Intrinsic religion was assessed with a single item measure drawn from Gorsuch and McPherson's (1989) psychometric work on data reduction. The item—*My whole life is based upon my religion*—is followed by a 5-point Likert-type response key, ranging from *strongly disagree* to *strongly agree*. This item was identified by Gorsuch and McPherson as the single best item to appropriately represent the complete nine-item intrinsic measure of religion.

Islamic practices were measured with the 10-item Islamic Behavioral Practices scale (Jana-Masri & Priester, 2007). Each item represents a common Islamic practice drawn from the Quran. A sample item is *I give Zakah (alms)*. Individuals indicate their response on a five-point response key, ranging from *never* to *always*. The Cronbach's alpha obtained in the original validation study was .81.

Life satisfaction was measured with a single item drawn from OECD Guidelines on Measuring Subjective Well-Being (OECD, 2013). This question asks, *In general, how satisfied are you with your life as a whole?* Individuals indicate their response on an 11-point response key that ranges from *completely dissatisfied* through to *completely satisfied*.

Depression was measured with a short form of the Center for Epidemiological Studies Depression Scale (CES-D; Melchior, Huba, Brown, & Reback, 1993). The four-item short form correlated .87 with the original 20 item CES-D. Individuals are asked how often they experienced depression-related feelings over the past week (e.g., *I felt depressed*). Individuals are presented with four options, ranging from *rarely/none* (less than 1 day) to *most or all of the time* (5–7 days). The Cronbach's alpha obtained in the original validation study was .81.

All measures were coded so that higher values indicate a greater amount of the construct measured. Thus, higher scores on, for example, the spirituality and depression scales indicate higher levels of spirituality and depressive symptoms.



The complete survey instrument was reviewed and discussed by the research team before administration. Minor changes were made to ensure cultural congruence with the intended study sample. For example, the original ISS defined spirituality in terms of a relationship with God or Ultimate Transcendence. In the present study this wording was changed so that spirituality was posited to refer to “your relationship with God.” These minor alterations are consistent with qualitative research on creating culturally congruent measures of spirituality and religion for use with Muslims (Berry et al., 2011). The Appendix includes the complete version of the ISS as used in the present study.

### Data Screening

Preliminary analyses were conducted to ascertain whether the data met the requirements for a CFA, a process that began with data cleaning. Two hundred eighty-five individuals completed most of the basic demographic questions.

Across the six items that comprise the ISS, missing data ranged from 4.9% to 6.0% of cases. Although no firm guidelines exist regarding what percentage of missing values is considered problematic, Kline (2011) suggests that 5% of missing cases is of little concern, particularly if the values are missing at random. This was the case in the present study as Little’s MCAR test indicated the data were missing completely at random (Kline, 2011). Missing data were imputed using the expectation–maximization (EM) algorithm procedure (Byrne, 2010).

To assess multivariate normality, Mardia’s (1980) test of multivariate normality was examined. Although the assumption of multivariate normality was not supported, maximum likelihood (ML) estimation is considered to be relatively robust (Schumacker & Lomax, 2010). Under conditions of moderate univariate non-normality (i.e., skew < 2, kurtosis < 7), with data consisting of at least five categories, Finney and DiStefano (2006) suggest the violation of multivariate normality can be considered relatively mild. Given that skew and kurtosis values fell within these criteria, the original, untransformed data were used.

A check for multivariate outliers was conducted. An examination of the Mahalanobis distance ( $D^2$ ) indicated four cases exerted a problematic influence (Byrne, 2010). These four cases were removed from the data set, resulting in a final sample of 281.

To obtain trustworthy estimates using ML estimation, at least five cases per model parameter are required (Mueller & Hancock, 2010). In the initial model in the present study, 12 distinct parameters were estimated. Thus, the sample size of 281 was judged adequate for data analysis.

### Data Analysis

The data were analyzed using AMOS 22.0 and SPSS 22.0. The CFA was conducted using the former and all other statistical procedures were conducted using the latter program. For the CFA the ML procedure was used to estimate the parameters. The indicator with the highest factor loading served as the reference item and was constrained to a value of 1.0 (Garson, 2012). Model evaluation included an examination of the model fit indices, modification indexes, standardized factor loadings, and squared multiple correlations.

Different fit indices can give different assessments of model fit, depending upon the type or class of index employed (Mueller &

Hancock, 2010). Accordingly, model fit was evaluated using the normed chi-square value, the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). These indices address the issue of fit from three different perspectives (i.e., incremental, parsimonious, and absolute) and are widely recommended (Byrne, 2010; Garson, 2012; Mueller & Hancock, 2010; Schumacker & Lomax, 2010; West, Taylor, & Wu, 2012). Fit was also assessed by examining the standardized residuals.

Cronbach’s alpha was computed to assess reliability, and criterion and concurrent validity were tested by computing correlation coefficients between the ISS and theoretically associated constructs. Before relating the results of these tests, however, descriptive statistics are reported.

## Results

### Descriptive Statistics

The sampling strategy produced a relatively diverse group of study participants (see Table 1). The typical participant was a 38-year-old, ( $SD = 11.72$ ) married female adherent of the Sunni tradition, which is by far the largest tradition within Islam. Most participants were highly educated as evidenced by the fact that 55% of respondents ( $n = 155$ ) had completed a graduate degree.

Just over half the sample was born outside the United States (52%,  $n = 147$ ) and the racial/ethnic composition of the sample reflected diverse nations of origin. The most common ethnicity was Middle Eastern (37%,  $n = 103$ ), followed closely by Asian (31%,  $n = 86$ ), then African American (14%,  $n = 39$ ), and European American (11%,  $n = 30$ ). Although 61% ( $n = 171$ ) generally spoke English at home, a substantial percentage spoke other languages.

To tease out some of the diversity in the sample, subgroup analyses were conducted based upon ethnicity, nation of origin, language spoken at home, and educational status. Among Middle Eastern respondents, 64% ( $n = 66$ ) were born outside the United States, a plurality primarily spoke Arabic at home (45%,  $n = 46$ ), and approximately 50% ( $n = 51$ ) had completed a graduate degree. Among Asians, 62% ( $n = 53$ ) were born outside the United States, 64% ( $n = 55$ ) spoke English at home, and 66% ( $n = 57$ ) had completed a graduate degree.

A somewhat different picture emerged for African Americans and European Americans. Among the former, 85% ( $n = 33$ ) were born in the United States, 95% ( $n = 37$ ) spoke English at home, and 56% ( $n = 22$ ) had completed a graduate degree. Among European Americans, 70% ( $n = 21$ ) were born in the United States, 80% ( $n = 24$ ) spoke English at home, and 50% ( $n = 15$ ) had completed a graduate degree.

### Confirmatory Factor Analysis

To test the hypothesis regarding model fit with self-identified Muslims, a measurement model was constructed. Spirituality served as the latent construct and the six items that comprise the ISS served as indicators or observed variables (see Figure 1). Overall model fit was assessed using the aforementioned fit indices—normed chi-square, CFI, RMSEA, and SRMR—in tandem with an examination of the standardized residuals.

Table 1  
Participant Characteristics (N = 281)

Characteristic	% <sup>1</sup>	N	M	SD
Age			38.18	11.72
Gender				
Female	69	193		
Male				
Religious tradition	31	87		
Sunni	85	239		
Shiite	5	14		
Other <sup>2</sup>	10	28		
(Just) Muslim	(8)	(22)		
Sufi/spiritual	(1)	(4)		
Agonistic/culturally Muslim	(1)	(2)		
Marital status				
Married	57	160		
Single	32	91		
Other <sup>2</sup>	11	30		
Divorced	(8)	(22)		
Engaged	(1)	(4)		
Separated	(1)	(3)		
Widowed	(0)	(1)		
Education				
Graduate degree	55	155		
4-year college degree	23	64		
Some college	15	43		
High school or less	7	19		
Nation of origin				
United States	48	134		
Other <sup>2</sup>	52	147		
Egypt	(9)	(25)		
Pakistan	(9)	(25)		
Bangladesh	(3)	(9)		
Canada	(2)	(7)		
Palestine	(2)	(6)		
Jordan	(2)	(5)		
Libya	(2)	(5)		
Iraq	(2)	(5)		
Saudi Arabia	(2)	(5)		
Syria	(2)	(5)		
Ethiopia	(1)	(4)		
Tunisia	(1)	(4)		
India	(1)	(3)		
Lebanon	(1)	(3)		
Kuwait	(1)	(2)		
Morocco	(1)	(2)		
Nigeria	(1)	(2)		
Turkey	(1)	(2)		
Other listed nations	(8)	(22)		
Unanswered	(2)	(6)		
Race/ethnicity				
Middle Eastern	37	103		
Asian	31	86		
African American/Black	14	39		
European American/White	11	30		
Hispanic/Latino	1	4		
Other <sup>2</sup>	7	19		
Mixed	(2)	(6)		
African	(1)	(3)		
Arab	(1)	(3)		
Muslim	(1)	(3)		
Unanswered	(1)	(4)		
Language spoken at home				
English	61	171		
Arabic	20	56		
Urdu	9	25		
Farsi	1	4		

(continued)

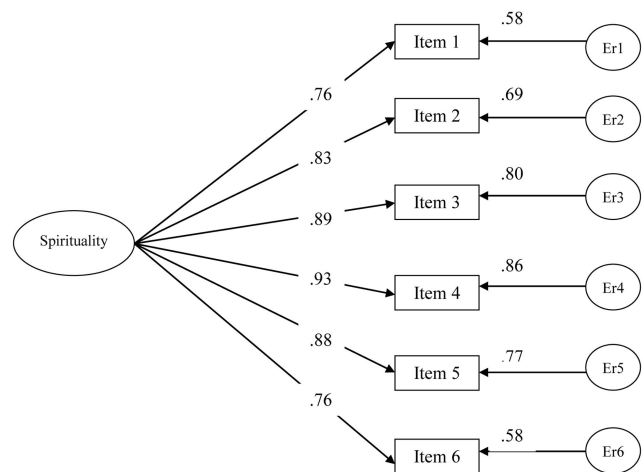
Characteristic	% <sup>1</sup>	N	M	SD
Other <sup>2</sup>	9	25		
Bengali	(1)	(4)		
Amharic	(1)	(2)		
French	(1)	(2)		
Turkish	(1)	(2)		
Other listed language	(3)	(8)		
Unanswered	(2)	(7)		

<sup>1</sup> Percentages may not equal 100 percent due to rounding. <sup>2</sup> This question included an additional option for indicating alternative responses.

The commonly reported chi-square test is sensitive to sample size (West et al., 2012). To address this problem, some SEM modelers use the normed chi-square. While a nonsignificant chi-square essentially indicates a perfect fit, a good fit is indicated when the ratio between the chi-square statistic and the degrees of freedom is less than 5 ( $\chi^2/df < 5$ ; West et al., 2012).

For the CFI, values  $> .90$  indicate a marginal fit between the proposed model and the data (Garson, 2012) while values  $> .95$  indicate a good fit (Byrne, 2010). For the RMSEA, values  $< .08$  represent a reasonable fit (Byrne, 2010), while values  $> .10$  indicate a poor fit (Garson, 2012). For the SRMR a value  $< .09$  represents a good fit (Mueller & Hancock, 2008). For the standardized residuals, values  $> 2.58$  suggest a poor fit between variables (Byrne, 2010).

The values for the initial model were as follows:  $\chi^2 = 35.43$  ( $df = 9, p < .001$ ), normed  $\chi^2 = 3.94$ , CFI = .98, RMSEA = .10 (90% CI [.07, .14]), and SRMR = .02. Take together, these values suggest a generally acceptable fit. The point estimate of the RMSEA is high and the confidence intervals are also wide, which may be an artifact of the non-normal distribution and the relatively small sample size (Kline, 2011). Aside from this one index, all values indicate a good fit between the proposed model and the data. An examination of the standardized residuals revealed that all values fell well below the 2.58 value, which is also consistent with the notion of an acceptable fit.



(table continues)

Figure 1. Measurement model with standardized parameter estimates.

An examination of the modification indexes suggested adding a measurement error covariance between Items 2 and 3 would improve the model fit. In contrast to exploratory factor analysis, some error terms may be allowed to correlate in a CFA. If, for example, the observed variables share similar language or tap a shared construct as occurred in this study, the addition of error covariance can be warranted (Blunch, 2008). Given that Items 2 and 3 tap a shared construct—namely intrinsic spirituality—the addition of an error covariance seemed justified.

The addition of the error covariance improved the model fit (not shown). The resulting values were as follows:  $\chi^2 = 20.03$  ( $df = 8$ ,  $p = .01$ ), normed  $\chi^2 = 2.50$ , CFI = .99, RMSEA = .07 (90% CI [.03, .11]), and SRMR = .02. These values represent a good to excellent fit across all indices.

The hypothesis regarding the validity and reliability of the individual indicators was assessed by examining the standardized factor loadings and squared multiple correlations ( $R^2$  values). In SEM, the factor loadings function as validity coefficients, indicating how accurately the item measures the latent construct (Schumacker & Lomax, 2010). Coefficients  $> .70$  indicate relatively high loadings (Kline, 2011).

The  $R^2$  values provide an assessment of the reliability of each observed measure with respect to its underlying latent construct (Blunch, 2008). Ideally, the  $R^2$  values should be  $> .50$  for each indicator in a CFA (Kline, 2011). Items that exhibit low validity or reliability should be removed or rephrased (Blunch, 2008).

As can be seen in Figure 1, the factor loadings ranged from .76 to .93. The  $R^2$  values ranged from .58 to .86. These values suggest that each observed variable functions as a valid and reliable indicator of spirituality with the study's sample.

Next Cronbach's alpha was examined. Consistent with expectations, internal consistency was high ( $\alpha = .93$ ). Sequential deletion of items yielded alphas ranging from .91 to .93.

Alpha coefficients and ISS scores were computed for each major ethnic group to assess possible differences. The obtained alphas were relatively consistent across the four major ethnic groups: Middle Eastern (.95), Asian (.92), African American (.89), and European American (.93), although it should be noted that sample size can affect alpha coefficients (Nunnally & Bernstein, 1994). The ISS score for the complete sample was 8.34 ( $SD = 1.55$ ). A one-way ANOVA using the Tukey post hoc test revealed no significant differences between ISS scores across the different ethnic groups.

### Criterion Validity

To test the hypothesis regarding criterion validity, Pearson's correlation coefficients were calculated between the ISS and the study's two religion measures: the intrinsic religion measure and the Islamic Behavioral Practices scale. The coefficient for the former was .66 and the coefficient for the latter was .59, values in keeping with the expected strong correlation between the ISS and these two religion measures.

These values are also consistent with the hypothesis that spirituality is related to, but distinct from, religion. It is important to note that other factors can cause a less than perfect correlation. Nevertheless, the coefficients obtained are consistent with the fact that ISS taps a construct that is not identical to religion.

### Concurrent Validity

To test the hypothesis regarding concurrent validity, Pearson's correlation coefficients were computed between the ISS and the life satisfaction and depression measures. The ISS was positively associated with life satisfaction,  $r = .32$ ,  $p < .01$ , and negatively associated with depression,  $r = -.13$ ,  $p < .05$ . The direction of both these associations is consistent with existing theory and research.

### Discussion

Muslims may be one of the fastest growing religious groups in the United States (Hedayat-Diba, 2014). Yet, despite the growing size of this culturally distinct population, few, if any, measures of spirituality have been validated with this population (Kapusinski & Masters, 2010; Monod et al., 2011). The purpose of this study was to address this gap in the literature by validating a spirituality measure for use with Muslims in the United States—specifically, the Intrinsic Spirituality Scale (ISS).

The ISS measures the degree to which respondents are motivated by their spirituality, defined as their relationship with God (or Ultimate Transcendence). Perhaps the central tenet of Islam is the belief that adherents should be motivated by their relationship with God (Koenig & Shohaib, 2014). This concept is embedded in the word Islam itself, which means submission to God. As such, the ISS measures a dimension of spirituality that is central to Islam, the notion that adherents should be guided, directed, or motivated by God in all aspects of their lives.

The results suggest the ISS is a valid and reliable measure of spirituality with Muslims. The measurement model fit the data well across fit indices (i.e., normed  $\chi^2$ , CFI, RMSEA, and SRMR) and each of the six indicators demonstrated acceptable levels of validity and reliability. The Cronbach's alpha coefficient obtained for the present sample was .93.

Criterion and concurrent validity were also evident. Regarding the former, the ISS correlated with the measures of intrinsic religion and Islamic practices in a strong manner that is consistent with the hypothesis that the ISS taps a construct that is related to, but distinct, from religion. Concurrent validity was demonstrated by the fact that the ISS correlated positively with life satisfaction and negatively with depression. All relationships, along with the measures used to operationalize the underlying constructs, are consistent with existing theory and research (Koenig et al., 2012), as well as common practice validating spirituality measures (Kapusinski & Masters, 2010; Monod et al., 2011).

### Implications for Practice

The results have important implications for clinical practice with Muslims. The use of measures that "make sense" within the context of an Islamic worldview is critical to effective service provision (Berry et al., 2011). For instance, the use of culturally sensitive assessment procedures helps build trust and therapeutic rapport (Sue & Sue, 2013).

Clinicians might use the ISS to assess clients' level of spiritual motivation (Pargament, 2007). The measure can be administered in a brief period of time and provides important insights into the degree to which spirituality may be a salient dimension of treatment. Individuals who score relatively high values might be can-

didates for a more extensive qualitative spiritual assessment and, potentially, spiritually based interventions (Nielsen, 2004).

A number of studies have examined the efficacy of therapeutic interventions that include a spiritual dimension (Tan, 2013). Perhaps most of the research has been conducted on cognitive-behavioral therapy that has been modified to incorporate clients' spiritual beliefs and practices (Hook et al., 2010). This modality has been classified as an empirically validated intervention for highly spiritual Christians dealing with depression, and shows promise as a treatment for similar disorders among members of other faith groups including Muslims (Hodge, 2006; Koenig & Shohaib, 2014).

Outcome studies with Muslims have demonstrated positive results using this modality to treat anxiety and depression (Azhar & Varma, 2000; Razali, Hasanah, Aminah, & Subramaniam, 1998; Razali, Aminah, & Khan, 2002). In addition, some evidence suggests that spiritually motivated Muslims may be more likely to benefit from such culturally accommodative therapies. The ISS can provide a quick and efficient screening method to identify clients that may be particularly likely to benefit from treatments that incorporate a spiritual or religious dimension.

### Directions for Future Research

Many avenues for future research exist. The concept of criterion validity is operationalized in different ways (Babbie, 2013). In the present study, criterion validity was operationalized as the degree to which the ISS correlated with two established measures of religion. Although this understanding is consistent with the existing research on spirituality measures (Monod et al., 2011), future researchers might examine the degree to which the ISS predicts certain hypothesized variables over time.

Validation is also needed with additional samples of Muslims. Ideally, nationally representative samples should be employed. In practice, however, this is often hard to achieve (Kapuscinski & Masters, 2010). Accordingly, additional psychometric work might be conducted with other samples of Muslims, both clinical and community based, to assess the validity of the ISS. Indeed, incorporating the ISS into studies along with other measures of health and well-being can, over time, shed light on the utility of the measure in different contexts (Abu-Raiya & Hill, 2014).

### Strengths and Limitations

The strengths and limitations of the present study should also be noted. Among the major strengths of the present study is the diversity of the sample. A significant concern regarding existing spirituality measures is the homogeneity of the samples used to validate the instruments (Kapuscinski & Masters, 2010; Monod et al., 2011). In addition to age, religion, class, and cultural background (e.g., young Christian students at a single university), validation samples often evidence minimal diversity in other areas as well. This is particularly problematic for spirituality measures as spirituality often varies based upon such characteristics (Newport, 2012).

To circumvent this problem, methods were employed in the current study to obtain a diverse sample. Within the broader Islamic tradition, the sample included different Islamic subtraditions, as well as substantial diversity in the areas of age, gender, marital status, nation of origin, race/ethnicity, language preference,

and to a lesser extent, education. Taken together, the diversity of the sample suggests the ISS may have utility with Muslims from many different backgrounds. This contention is supported by the fact that the Cronbach's alphas and ISS scores obtained were relatively consistent across the major ethnic groups in the study.

Although the array of Muslim organizations sampled produced a relatively diverse group of study participants, the sampling strategy also limits the generalizability of the results. The nonrandom sampling procedure precludes generalizing the results beyond the present sample. For instance, it is plausible that more committed, highly educated Muslims disproportionately participated in the study. As a result, the measure may have less utility with less educated Muslims. It should be noted, however, that Muslims are more likely to have graduated from college than most Americans (Hedayat-Diba, 2014). Although additional psychometric work is clearly warranted, the present study does establish some preliminary validity for the ISS with Muslims.

### Conclusion

The United States is growing increasingly diverse in multiple ways, including religiously (Eck, 2001). One of the more frequent criticisms of existing spirituality measures is the lack of instruments that have been validated with non-Christian populations (Berry et al., 2011). This study responds to this criticism by validating a spirituality measure for use with the nation's growing Muslim population.

Based upon the CFA and tests of criterion and concurrent validity with the diverse sample used in this study, the ISS appears to be a valid and reliable measure of spirituality for Muslims in the United States and perhaps other nations as well. The six-item ISS can be administered quickly and provides insight into the extent to which spirituality functions as a motivational drive in clients' lives. Although more work remains to be done, this study represents one small step forward in the journey toward more culturally competent service provision.

### References

- Abu-Raiya, H., & Hill, P. C. (2014). Appraising the state of measurement of Islamic religiousness. *Psychology of Religion and Spirituality, 6*, 22–32. <http://dx.doi.org/10.1037/a0035082>
- Allport, G. W., & Ross, J. M. (1967). Personal religious orientation and prejudice. *Journal of Personality and Social Psychology, 5*, 432–443. <http://dx.doi.org/10.1037/h0021212>
- Azhar, M. Z., & Varma, S. L. (2000). Mental illness and its treatment in Malaysia. In I. Al-Issa (Ed.), *Al-Junun: Mental illness in the Islamic world* (pp. 163–185). Madison, CT: International Universities Press.
- Babbie, E. (2013). *The practice of social research (13th)*. Belmont, CA: Wadsworth.
- Bagby, I. (2011). *The American Mosque 2011. Council on American-Islamic relations*. Retrieved from <http://faithcommunitiestoday.org/sites/faithcommunitiestoday.org/files/The%20American%20Mosque%202011%20web.pdf>
- Barrios, M., Villarroja, A., Borrego, A., & Olle, C. (2011). Response rates and data quality in web and mail surveys administered to PhD holders. *Social Science Computer Review, 29*, 208–220. <http://dx.doi.org/10.1177/0894439310368031>
- Berry, D. M., Bass, C. P., Forawi, W., Neuman, M., & Abdallah, N. (2011). Measuring religiosity/spirituality in diverse religious groups: A consideration of methods. *Journal of Religion and Health, 50*, 841–851. <http://dx.doi.org/10.1007/s10943-011-9457-9>



- Blunch, N. J. (2008). *Introduction to structural equation modeling using SPSS and AMOS*. Thousand Oaks, CA: Sage.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications and programming* (2nd ed.). New York, NY: Taylor and Francis.
- de Klerk, J. J., Boshoff, A. B., & Wyk, R. (2009). Measuring the meaning of life in South Africa: Validation of an instrument developed in the USA. *South African Journal of Psychology, 39*, 314–325. <http://dx.doi.org/10.1177/008124630903900306>
- Dwairy, M. (2006). *Counseling and psychotherapy with Arabs and Muslims: A culturally sensitive approach*. New York, NY: Teachers College Press.
- Eck, D. L. (2001). *A new religious America*. New York, NY: HarperCollins.
- Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Greenwich, CT: Information Age Publishing.
- Garson, G. D. (2012). *Structural equation modeling*. Asheboro, NC: Statistical Publishing Associates.
- Ghorbani, N., Watson, P. J., & Khan, Z. H. (2007). Theoretical, empirical, and potential ideological dimensions of using Western conceptualizations to measure Muslim religious commitments. *The Journal of Muslim Mental Health, 2*, 113–131. <http://dx.doi.org/10.1080/15564900701613041>
- Gorsuch, R. L., & McPherson, S. E. (1989). Intrinsic/Extrinsic measurement: I/E-revised and single item scales. *Journal for the Scientific Study of Religion, 28*, 348–354. <http://dx.doi.org/10.2307/1386745>
- Gough, H. R., Wilks, S. E., & Prattini, R. J. (2010). Spirituality among Alzheimer's caregivers: Psychometric reevaluation of the Intrinsic Spirituality Scale. *Journal of Social Service Research, 36*, 278–288. <http://dx.doi.org/10.1080/01488376.2010.493848>
- Graham, J. R., Bradshaw, C., & Trew, J. L. (2010). Cultural considerations for social service agencies working with Muslim clients. *Social Work, 55*, 337–346. <http://dx.doi.org/10.1093/sw/55.4.337>
- Hall, R. E. (2012). Islamic spirituality vis-a-vis Asian Pacific Muslim populations: A resource for Western social work practice. *International Social Work, 55*, 109–124. <http://dx.doi.org/10.1177/0020872811409472>
- Hedayat-Diba, Z. (2014). Psychotherapy with Muslims. In P. S. Richards & A. E. Bergin (Eds.), *Handbook of psychotherapy and religious diversity* (2nd ed., pp. 287–317). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/14371-012>
- Hill, P. C., & Hood, R. W. (Eds.). (1999). *Measures of religiosity*. Birmingham, AL: Religious Education Press.
- Hodge, D. R. (2003). The Intrinsic Spirituality Scale: A new six-item instrument for assessing the salience of spirituality as a motivational construct. *Journal of Social Service Research, 30*, 41–61. [http://dx.doi.org/10.1300/J079v30n01\\_03](http://dx.doi.org/10.1300/J079v30n01_03)
- Hodge, D. R. (2006). Spiritually modified cognitive therapy: A review of the literature. *Social Work, 51*, 157–166. <http://dx.doi.org/10.1093/sw/51.2.157>
- Hodge, D. R., & Gillespie, D. F. (2005). Phrase completion scales. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 3, pp. 53–62). San Diego, CA: Academic Press. <http://dx.doi.org/10.1016/B0-12-369398-5/00124-9>
- Hook, J. N., Worthington, E. L., Jr., Davis, D. E., Jennings, D. J., II, Gartner, A. L., & Hook, J. P. (2010). Empirically supported religious and spiritual therapies. *Journal of Clinical Psychology, 66*, 46–72.
- Husain, A., & Ross-Sheriff, F. (2011). Cultural competence with Muslim Americans. In D. Lum (Ed.), *Culturally competent practice: A framework for understanding diverse groups and justice issues* (4th ed., pp. 358–389). Belmont, CA: Brooks/Cole.
- Jana-Masri, A., & Priester, P. E. (2007). The development and validation of a Qur'an-based instrument to assess Islamic religiosity: The religiosity of Islam scale. *Journal of Muslim Mental Health, 2*, 177–188. <http://dx.doi.org/10.1080/15564900701624436>
- Ji, C.-H. C., & Ibrahim, Y. (2007). Islamic doctoral orthodoxy and religious orientations: Scale development and validation. *International Journal for the Psychology of Religion, 17*, 189–208. <http://dx.doi.org/10.1080/10508610701402192>
- Kapuscinski, A. N., & Masters, K. S. (2010). The current status of measures of spirituality: A critical review of scale development. *Psychology of Religion and Spirituality, 2*, 191–205. <http://dx.doi.org/10.1037/a0020498>
- Kline, R. B. (2011). *Principles and practices of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Kobeisy, A. N. (2004). *Counseling American Muslims*. Westport, CT: Praeger.
- Koenig, H. G. (Ed.). (1998). *Handbook of religion and mental health*. San Diego, CA: Academic Press.
- Koenig, H. G., King, D., & Carson, V. B. (2012). *Handbook of religion and health* (2nd ed.). New York, NY: Oxford University Press.
- Koenig, H. G., & Shohaib, S. A. (2014). *Health and well-being in Islamic societies: Background, research, and applications*. New York, NY: Springer. <http://dx.doi.org/10.1007/978-3-319-05873-3>
- Mardia, K. V. (1980). Tests of univariate and multivariate normality. In P. R. Krishnaiah (Ed.), *Handbook of statistics, Vol. 1: Analysis of variance* (pp. 279–320). Amsterdam, The Netherlands: Nonh-Holland.
- Melchior, L. A., Huba, G., Brown, V. B., & Reback, C. J. (1993). A short depression index for women. *Educational and Psychological Measurement, 53*, 1117–1125. <http://dx.doi.org/10.1177/0013164493053004024>
- Monod, S., Brennan, M., Rochat, E., Martin, E., Rochat, S., & Büla, C. J. (2011). Instruments measuring spirituality in clinical research: A systematic review. *Journal of General Internal Medicine, 26*, 1345–1357. <http://dx.doi.org/10.1007/s11606-011-1769-7>
- Mueller, R. O., & Hancock, G. R. (2008). Best practices in structural equation modeling. In J. Osborne (Ed.), *Best practices in quantitative methods* (pp. 488–508). Thousand Oaks, CA: Sage. <http://dx.doi.org/10.4135/9781412995627.d38>
- Mueller, R. O., & Hancock, G. R. (2010). Structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods* (pp. 371–383). New York, NY: Taylor and Francis.
- Newport, F. (2012). *God is alive and well: The future of religion in America*. New York, NY: Gallup Press.
- Nielsen, S. L. (2004). A Mormon rational emotive behavior therapist attempts Qur'anic rational emotive behavior therapy. In P. S. Richards & A. E. Bergin (Eds.), *Casebook for a spiritual strategy in counseling and psychotherapy* (pp. 213–230). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10652-013>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- OECD. (2013). *OECD guidelines on measuring subjective well-being*. Retrieved from <http://www.oecd.org/statistics/Guidelines%20on%20Measuring%20Subjective%20Well-being.pdf>
- Pargament, K. I. (2007). *Spiritually integrated psychotherapy: Understanding and addressing the sacred*. New York, NY: Guilford Press.
- Pargament, K. I. (2013). *APA Handbook of psychology, religion, and spirituality*. Washington, DC: American Psychological Association.
- Pew Research Center. (2011). *Muslim Americans: No signs of growth in alienation or support for extremism*. Retrieved from <http://www.people-press.org/2011/08/30/muslim-americans-no-signs-of-growth-in-alienation-or-support-for-extremism/>
- Razali, S. M., Aminah, K., & Khan, U. A. (2002). Religious-cultural psychotherapy in the management of anxiety patients. *Transcultural Psychiatry, 39*, 130–136. <http://dx.doi.org/10.1177/136346150203900106>
- Razali, S. M., Hasanah, C. I., Aminah, K., & Subramaniam, M. (1998). Religious–sociocultural psychotherapy in patients with anxiety and de-

pression. *Australian and New Zealand Journal of Psychiatry*, 32, 867–872. <http://dx.doi.org/10.3109/00048679809073877>

Richards, P. S., & Bergin, A. E. (Eds.). (2014). *Handbook of psychotherapy and religious diversity* (2nd ed.). Washington, DC: American Psychological Association.

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Taylor and Francis.

Shin, E., Johnson, T. P., & Rao, K. (2012). Survey mode effects on data quality: Comparison of web and mail modes in a U.S. national panel survey. *Social Science Computer Review*, 30, 212–228. <http://dx.doi.org/10.1177/0894439311404508>

Smith, C. (2003). *The secular revolution*. Berkeley, CA: University of California Press.

Sue, D., & Sue, D. (2013). *Counseling the culturally diverse: Theory and practice* (6th ed.). Hoboken, NJ: Wiley.

Tan, S.-Y. (2013). Addressing religion and spirituality from a cognitive-behavioral perspective. In K. I. Pargament (Ed.), *APA handbook of psychology, religion, and spirituality: Vol. 2. An applied psychology of religion and spirituality* (pp. 169–187). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/14046-008>

Van Hook, M., Hugen, B., & Aguilar, M. A. (Eds.). (2001). *Spirituality within religious traditions in social work practice*. Pacific Grove, CA: Brooks/Cole.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York, NY: Guilford Press.

Zinnbauer, B. J., Pargament, K. I., Cole, B., Rye, M. S., Butter, E. M., Belavich, T. G., . . . Kadar, J. L. (1997). Religion and spirituality: Unfuzzifying the fuzzy. *Journal for the Scientific Study of Religion*, 36, 549–564. <http://dx.doi.org/10.2307/1387689>

### Appendix

#### Intrinsic Spirituality Scale for Muslims

The following six questions use a 0-to-10 response key in which 0 corresponds to an absence or zero amount of the attribute, while 10 corresponds to the maximum amount of the attribute. Please **circle the number** along the continuum that best reflects your initial feeling. Spirituality refers to your relationship with God.

1. In terms of the questions I have about life, my spirituality answers <b>no questions</b>	<u>0</u>	1	2	3	4	5	6	7	8	9	<b>absolutely all my questions</b> <u>10</u>
2. Growing spiritually is <b>more important than anything else in my life</b>	<u>10</u>	9	8	7	6	5	4	3	2	1	<b>of no importance to me</b> <u>0</u>
3. When I am faced with an important decision, my spirituality <b>plays absolutely no role</b>	<u>0</u>	1	2	3	4	5	6	7	8	9	<b>is always the overriding consideration</b> <u>10</u>
4. Spirituality is <b>the master motive of my life, directing every other aspect of my life</b>	<u>10</u>	9	8	7	6	5	4	3	2	1	<b>not part of my life</b> <u>0</u>
5. When I think of the things that help me to grow and mature as a person, my spirituality <b>has no effect on my personal growth</b>	<u>0</u>	1	2	3	4	5	6	7	8	9	<b>is absolutely the most important factor in my personal growth</b> <u>10</u>
6. My spiritual beliefs affect <b>absolutely every aspect of my life</b>	<u>10</u>	9	8	7	6	5	4	3	2	1	<b>no aspect of my life</b> <u>0</u>

Note. Adapted with permission from Taylor & Francis from Hodge (2003). Available from [http://www.tandfonline.com/doi/abs/10.1300/J079v30n01\\_03#](http://www.tandfonline.com/doi/abs/10.1300/J079v30n01_03#)

Received August 25, 2014

Revision received February 28, 2015

Accepted March 3, 2015 ■